GridNet: efficiently learning deep hierarchical representation for 3D point cloud understanding

Huiqun WANG, Di HUANG, Yunhong WANG

Frontiers of Computer Science, DOI: 10.1007/s11704-020-9521-2

Problems & Ideas

- The data structurization in point cloud understanding is not efficient enough
  - Point-based: Hard to describe detailed spatial point distribution.
  - Voxel-based: Always incur high computation and memory costs.
- Ideas: Combining the advantages of holistic description of voxel-based methods and detailed description of point-based ones.
  - Introducing virtual anchors for global description.
  - Employing sampled points for local description.

Main Contributions

- **Contributions**
  - **We propose GridNet, a lightweight and efficient deep network with a novel data structurization method for 3d point cloud representation**
  - **We extend GridNet to a framework of Grid-Encoder and Grid-Decoder for segmentation**
  - **We show competitive results in both classification and segmentation tasks.**

- **Time and space complexity**

	PointNet++	PointCNN	GridNet	GridNet-S
Forwarding Time	~164ms	~93ms	~62ms	~22ms
Params	1.48M	0.6M	1.6M	0.6M
Accuracy	90.70%	92.20%	92.20%	91.71%
Converging Time	13 hours	18 hours	6 hours	4 hours

- **Classification Accuracy on ModelNet 40**

Input	Method	Model Size	Accuracy
Volumetric	3D ShapeNets	38M	77.32
Volumetric	VRN	18M	91.33
Volumetric	VRN-ensemble	90M	95.54
MultiView+Vol	FusionNet	118M	90.80
MultiView+Image	MV-CNN	VGG-based	90.10
MultiView+PC	PVNet	AlexNet-based	93.20
PC+Vol	PointGrid	5M	92.00
OCT-Tree	OCNN	-	90.60
3D Distance Field	FPNN	-	97.50
Pointwise MLP	PointNet	3.48M	89.20
Pointwise MLP	PointNet++(XYZ)	1.48M	90.70
X-Conv	PointCNN(XYZ)	0.60M	92.20
Grid-Encoder	GridNet(XYZ)	1.60M	91.90
Pointwise MLP	PointNet++(XYZ+N)	1.48M	91.90
X-Conv	PointCNN(XYZ+N)	0.60M	92.61
Grid-Encoder	GridNet(XYZ+N)	1.60M	92.70